R E P O R T R E S U M E S

ED 016 255 48 CG 001 327

INTELLIGENCE TESTS AND EDUCATIONALLY-RELEVANT MEASUREMENTS.

BY- BARRITT, LOREN S.
MICHIGAN UNIV., ANN ARBOR,CTR.FOR RES.LANG.AND BEH

REPORT NUMBER BR-6-1784 PUB DATE 1 SEP 67
CONTRACT OEC-3-6-061784-0508
EDRS PRICE MF-$0.25 HC-$0.48 10P.

DESCRIPTORS- *INTELLIGENCE TESTS, INTELLIGENCE, *TEST VALIDITY, *TESTING PROBLEMS, MEASUREMENT GOALS,

THE RELEVANCE OF INTELLIGENCE TESTS FOR EDUCATIONAL USES IS CHALLENGED ON TWO GROUNDS--(1) TESTS WHICH MERELY PREDICT THE LIKELIHOOD OF FUTURE SUCCESS DO NOT PROVIDE USEFUL INFORMATION FOR THOSE WHO WISH TO PRESCRIBE TREATMENTS TO ENHANCE PERFORMANCE, AND (2) INTELLIGENCE IS NOT DEFINED AND HENCE THE INTERPRETATION OF SCORES IS MISLEADING. IT IS SUGGESTED THAT EDUCATIONALLY-RELEVANT TESTS MUST BE BASED UPON EXPERIMENTAL PROCEDURES FOR VALIDATION. THE SKILLS TO BE TESTED MUST BE DEFINED AND REPORTED CLEARLY. THE MUTUAL BENEFITS OF SUCH REVISED PROCEDURES FOR PSYCHOLOGY AND EDUCATION ARE SUGGESTED. THIS REPORT IS INCLUDED IN "STUDIES IN LANGUAGE AND LANGUAGE BEHAVIOR," PROGRESS REPORT 5, SEPTEMBER 1, 1967. (AUTHOR)

ED016255

BR 6-1784
PA 48

Intelligence Tests and Educationally-Relevant Measurements[1, 2]

Loren S. Barritt
Center for Research on Language and Language Behavior
University of Michigan

The relevance of intelligence tests for educational uses is challenged on 2 grounds: (1) Tests which merely predict the likelihood of future success do not provide useful information for those who wish to prescribe treatments to enhance performance, (2) Intelligence is not defined and hence the interpretation of scores is misleading.

It is suggested that educationally-relevant tests must be based upon experimental procedures for validation. The skills to be tested must be defined and reported clearly. The mutual benefits of such revised procedures for psychology and education are suggested.

The concept of general intelligence and the instruments which provide its operational definition have been with us for a long time. Recently, however, a concern about the use of intelligence tests with minority group children has caused at least one large school system (New York City) to remove these tests from use (Loretan, 1966). I do not propose to deal with the specific issues raised by "civil rights" groups about these tests but prefer, rather, to deal with the broader questions which ought to be asked about the utility of intelligence tests in school settings.

McNemar addressed himself to these broader issues recently in his presidential address to the APA (McNemar, 1964). He began by asking the rhetorical question, "Lost our intelligence? Why?" The double entendre was, I am sure, intentional. McNemar defended intelligence tests against the onslaught of civil rights groups, "legislators," "school people" and statistically naive researchers. Any attempt on my part to interpret his remarks must of necessity lose something in translation. His argument seems to boil down to this:

1. General intelligence tests are able to predict school achievement as well as, or better than, anything else available, including differential aptitude batteries, and measures of creativity. They are, therefore, useful "tools."

2. Because they predict so well, we ought not to abandon intelligence now that it is being attacked. If we do, we might wind up with something which is not as "socially useful."

The above statements are generally accepted by psychologists and educators who deal with these instruments. I find myself in disagreement with this point of view.

CG 001 327

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

It is my conviction that intelligence tests are not useful tools in an educational setting and, further, that there is little danger in abandoning the tests because we can do much better. It is the purpose of this paper to explicate the reasons why intelligence is not a useful educational concept, and to suggest better alternatives.

McNemar's paper is a useful point of departure for an examination of the issues because it is the work of a highly-respected scholar in the field who has set forth the generally-accepted reasons for the tests. His paper is a useful foil for this discussion.

Intelligence Tests are Good Predictors

Certainly it must be acknowledged that intelligence tests are capable of predicting success in school, as well as, or better than, other psychological tests. The ability to predict does not by itself, however, guarantee the value of an instrument. This issue is a critical one because it is generally accepted that the demonstration of correlation with some relevant criterion is sufficient evidence for the value of a test. I would dispute such a claim.

My disagreement is based on the conviction that the scientist, who works on educational problems and who claims social utility for his procedures, cannot divorce himself from normative questions. Education is a value-loaded enterprise, and all those who investigate it participate in making value decisions. The point is well-illustrated in the present case. It is necessary to consider the goal of any test before deciding what evidence is required to demonstrate its value or utility. It seems clear that the goals of education are not adequately served by a test which does no more than predict future success.

It is the purpose of education to change behavior by providing the most succinct, effective treatments possible. To do this it is necessary to have information about intra- and inter-individual differences, and their interaction with various learning programs. A test which only predicts the likelihood of future success does not provide this information.

Let me illustrate this point with an analogy to the medical profession. The medical model is the right one in the present case because the goals, diagnosis and treatment, are congruent with those in education. Suppose a doctor were to give you a test of general health. If he said, "You have a low score; your chances for life are not very good," such a prediction might be extremely accurate but most of us would find the information useless because it does not identify the areas of difficulty and hence suggests no program of treatment for

remediation of the problem. Fortunately, doctors do not operate with a single test but, rather, use several tests which examine specific areas of functioning. A diagnosis is made based upon this information. A program of remediation can then be built upon this differential diagnosis. This is the procedure which we should seek to emulate in an educational setting.

> Prediction by itself is not helpful. As Cronback has said:
> Predicting outcome has no social value unless the psychologist or the subject himself can use the information to make better choices of treatment. The prediction must help to determine a treatment for every individual [1958].

I. Q. Interpretation

The I. Q. which an intelligence test yields not only fails to provide information about diagnosis and treatment, but there is evidence (Rosenthal & Jacobson, 1966) that it can effect the behavior of teachers toward children without their being aware of it. One might say that this is the fault of the teacher who does not know how to use the score, which is, in one sense, true. Teachers are not technically sophisticated with respect to test interpretation. But what could the technically sophisticated say about the I. Q. that would be different from the teacher's reaction? What is the appropriate interpretation of an I. Q. score?

The answer to these questions must rest upon the definition or meaning of intelligence. The construct of intelligence has proven rather difficult to define and it has even become acceptable to deny the need for definition (McNemar, 1964; Hayes, 1962). Most, perhaps all, of the difficulties discussed in this paper can be traced to this lack of definition.

Measurement is the process of assigning numerals to represent properties (Campbell, 1920). It is a shorthand which permits the manipulation of properties that would otherwise be inaccessible. Though psychologists have been unable to clarify what is meant by intelligence, it has nevertheless been measured. How do we know it is intelligence that has been measured? It is because our measurements correlate with other intellectual attainments. The circularity is well known.

What does not seem to be recognized, however, is that as a result of this lack of clarity in definition, teachers, when asked to translate the shorthand back into behavior, provide their own definition for intelligence. This is done implicitly, but done, nonetheless. These definitions all carry to one degree or another the idea of inherited ability which, at least in some measure, is a part of most psychological definitions of intelligence.

The fixity of intelligence is built into the I.Q. and contributes to misinterpretation. What can a teacher do to change nature or the I.Q.?

It is popular today to denigrate the view that intelligence is primarily an inherited characteristic. This viewpoint is not reflected in the I.Q. scores themselves, which are still stable over time (by design) and, consequently, resistant to change with learning. It is difficult for anyone technically sophisticated, or not, to interpret an I.Q. as anything but a score, fixed, in part, by inheritance, which identifies bright, average and dull children. In an educational setting, where our aim is to expand the intellect, it would seem beneficial to discontinue the use of tests which are misleading in this way.

Even if one believes in the primacy of nature over nurture and the fixity of intellectual development, there is cause to question whether this conception is a useful one for an <u>educational</u> program. Education ought to be guided by a search for effective treatments. Until there is clear evidence that we <u>cannot</u> teach certain children to amplify their intellectual skills (and when we get this much information we will probably have the ability to alleviate any such handicaps), we ought to continue to try. It might be said that education seeks to overcome nature with nurture. This may be an impossible task, but if one strives toward this end there is no need to fear that we have not aimed high enough.

<u>Educationally Relevant Measurements</u>

The utility of intelligence tests has been challenged here on two grounds. First, that the prediction of future success is not useful in an educational setting unless there is some prescription for treatment provided and, second, that scores are inherently misinterpretable and lead teachers, and others, to unfortunate conclusions about children.

It is not untrue to say that intelligence tests provide teachers with answers to questions which were never asked. Our technical skills seem to lead us astray. Obviously in such cases the tail is wagging the dog.

Education asks about the relationship between treatments and outcomes. What it seeks is information which will help to maximize treatments to enhance outcomes. These are questions about causes and effects. Correlational studies do not provide evidence to resolve such questions. What is required are experimental studies of the relationships between student characteristics, educational treatments, and outcomes. These studies might well use information from

correlational research to build hypotheses about these relationships. Intelligence tests might provide a reservoir of skills to be measured in conjunction with treatments to determine the nature of the interactions.

The development of treatment-relevant tests should be accompanied by a change in the procedure for reporting test scores to teachers. At present, scores are reported with reference to some group of people who are supposed to be relatively similar to the students we are testing. This assumption is usually a tenuous one. Further, this reference point has little value in the development of treatment programs. The reason for this in the case of the I.Q. is clear. The property we have tried to measure is not represented by the system we have chosen for the translation of raw scores. We have related scores to people. What we have is a kind of sociological yardstick. What we need is a system for reporting scores that is related to the treatments which will be prescribed. We need norms which relate scores to skills, rather than to some nebulous mean of a large group.

The principle involved in the translation of raw scores by reference to averages for groups is directed at specifying what should be expected of children at some point in their life. It tries to represent the individual performance but ends up describing a group. The value of studying groups to describe individuals seems limited at best. The specifying of the group mean as a point of reference does not accomplish the goal of representing individual performance norms.

To relate scores to skills would certainly require a more rigorous delineation of that which needs to be mastered and taught. It would further require precise definitions of skills so that meaningful numbers could be derived.

In sum, what I am suggesting is a change in the procedures for validating educational test scores. Since we want to know about the relationship between treatments and outcomes, we must use procedures that speak to such questions, i.e., experimental procedures. Further scores should be reported with reference to the skills, rather than the people, tested. The development of skill-norms is possible with the changes in procedure suggested above.

Language and Language Behavior

Language facility is the coin of the realm in schools. You either have it or you fail. Most intelligence tests place heavy emphasis upon language skills. It seems reasonable to expect that a good beginning toward the goal of educationally-relevant assessment will be attained by measurement of language variables.

It should be clear that if and when we begin to measure those aspects of language performance that are related to educational treatments we will at the same time be improving our understanding of human cognition. It should also be clear that by using experimental tryout procedures in the classroom we can begin to take insights directly from the laboratory and test their value in classrooms. This is precisely the kind of interaction between laboratory research and educational treatments that the Center for Research on Language and Language Behavior can make, and has already made, possible.

Education attempts to teach skills which can amplify the human mind (Bruner, 1965). Psychology has developed procedures for the study of human behavior. There is virtue in the combination of these enterprises. Education with its goals can direct psychology with its procedures to more relevant and useful research. A closer cooperation is to the benefit of both groups. Psychology has allowed itself to lose sight of the fascinating problems of human learning and performance. Education has often chosen the most expedient non-solution for its considerable problems. A wedding of the two areas should make it easier to justify research into relevant problems with appropriate techniques. A center for research on language and language behavior is a natural place for the ceremony to take place.

Footnotes

[1]This research reported herein was performed pursuant to Contract OEC-3-6-061784-0508 with the U.S. Department of Health, Education and Welfare, Office of Education, under the provisions of P.L. 83-531, Cooperative Research, and the provisions of Title VI, P.L. 85-864, as amended. This research report is one of several which have been submitted to the Office of Education as *Studies in language and language behavior, Progress Report V,* September 1, 1967.

[2]This article will appear in E.M. Zale (Ed.), Proceedings of the Conference on Language and Language Behavior. New York: Appleton-Century-Crofts, in press.

References

Anastasi, Anne. Psychology, psychologists and psychological testing. *American Psychologist*, 1967, *22*, 297-306.

Bruner, Jerome S. The growth of mind. *American Psychologist*, 1965, *20*, 1007-1017.

Burt, Cyril. The evidence for the concept of intelligence. *British Journal of Educational Psychology*, 1955, *25*, 158-177.

Campbell, N. R. *Physics, the elements*. London: Cambridge University Press, 1920.

Carroll, John B. Factors of verbal achievement. *Proceedings of the 1961 Invitational Conference on Testing Problems*. Princeton, N.J.; Educational Testing Service, 1962. Pp. 11-18.

Carroll, John B. The nature of the data, or how to choose a correlation coefficient. *Psychometrika*, 1961, *26*, No. 4, 347-372.

Combs, Arthur W. Intelligence from a perceptual point of view. *Journal of Abnormal and Social Psychology*, 1952, *47*, 662-673.

Cronbach, Lee J. The two disciples of scientific psychology. *American Psychologist*, 1957, *12*, 671-684.

Ferguson, George A. On learning and human ability. *Canadian Journal of Psychology*, 1954, *8*, 95-112.

Ferguson, George A. On transfer and the abilities of man. *Canadian Journal of Psychology*, 1956, *10*, 121-131.

Guilford, J. P. Three faces of intellect. *American Psychologist*, 1959, *14*, 469-479. (Republished in R. C. Anderson & D. P. Ausubel (Eds.), *Readings in the psychology of cognition*, 1965. Pp. 194-214.)

Hayes, Keith J. Genes, drives and intellect. *Psychological Reports*, 1962, *10*, 299-342.

Loretan, J. O. Alternatives to intelligence testing. *Proceedings of the 1965 Invitational Conference on Testing Problems*. Princeton, N.J.: Educational Testing Service, 1966. Pp. 19-30.

McNemar, Quinn. Lost: our intelligence? Why? *American Psychologist*, 1964, *12*, 871-882.

Rosenthal, R., & Jacobson, Lenore. Teachers' expectancies: determinants of pupils' I.Q. gains. *Psychological Reports*, 1966, *19*, 115-118.

NOV 21 1967

Sigel, Irving E. How intelligence tests limit understanding of intelligence. *Merrill-Palmer Quarterly*, 1963, *9*, 39-56.

Vernon, Phillip E. Ability factors and environmental influences. *American Psychologist*, 1965, *20*, 723-733.

Whiteman, Martin. Intelligence and learning. *Merrill-Palmer Quarterly*, 1964, *10*, 297-309.